

## **Predicting the Airbnb House Price**

Nowadays, it is really popular for people to book rooms on Airbnb when going to another city. However, there are many features of a property that may affect the price. We are interested in how these features affect the pricing on the Airbnb property. And it is a challenging problem that requires efforts and time to solve.

In this problem, we use a dataset provided by Kaggle, which includes a full list of properties posted on Airbnb, together with information on bathrooms, bedrooms, property types, etc. We want to predict the house price, given all the features of a property. To tackle this problem, we plan to do data preprocessing, then apply basic machine learning and tools for analytics techniques we learned in class.

Our original project proposal mentioned that we were supposed to predict the house price by analyzing the log-error between the estimate and the actual sale price. The original dataset we plan to analyze only have 1421 rows, we turn to analyze a larger dataset which has 65282 rows and aim to predict the `log_price` for Airbnb houses. The whole logic, analytics methods and data source are quite similar but the result is clearer and more meaningful than the previous one.

Last but not the least, we promise not to copy or plagiarize the code on Kaggle.

Group members: Yusang Liu, Yuhui Gao, Xuchen Qiu, Yashi Huang